Identification of an optimal Ncx binding sequence required for transcriptional activation

Hironori Shimizu[a,b], Myengmo Kang[a], Yoshinori Iitsuka[a], Masaharu Ichinose[b], Takeshi Tokuhisa[a], Masahiko Hatano[a,*]

[a] *Department of Developmental Genetics, Chiba University Graduate School of Medicine, Chiba 260-8670, Japan*
[b] *Department of Plastic and Reconstructive Surgery, Chiba University School of Medicine, Chiba 260-8670, Japan*

Received 24 February 2000; received in revised form 15 May 2000

Edited by Masayuki Miyasaka

Abstract **The *Ncx* gene encodes a homeobox containing transcription factor that belongs to the *Hox11* gene family. We determined specific Ncx protein binding consensus DNA sequences. Optimal Ncx binding sequences were 5′-CGGTAATTGG-3′ (TAAT core) and 5′-CGGTAAGTGG-3′ (TAAG core), which coincided with the Hox11 binding sequence. Both Ncx and Hox11 could bind to the TAAT and the TAAG core oligonucleotide in vitro. However, they could efficiently transactivate the reporter plasmid linked to the TAAT core sequence but not to the TAAG core sequence. Thus, Ncx and Hox11 act as transcriptional activators via their target sequence, 5′-CGGTAATTGG-3′.** **© 2000 Federation of European Biochemical Societies. Published by Elsevier Science B.V. All rights reserved.**

Key words: Ncx; Hox11; DNA-binding site; Homeobox protein

1. Introduction

Mammalian homeobox proteins have a conserved 60 amino acid sequence (homeodomain), which bind to a specific DNA sequence and function as a transcription factor [1–3]. They play an important role in specification of position along axis in the developing embryo [4] and in determination of cell type specification [5]. Disregulated expression of some of these proteins is responsible for tumorigenesis and congenital anomalies in humans and mice [6,7]. The *Hox11* gene was originally isolated from the chromosomal translocation breakpoint of T cell leukemia with t(10, 14) or t(7, 10) [8–11]. Hox11 is normally expressed in primordium of spleen, branchial arches, and some cranial ganglia such as trigeminus and facial ganglia [12,13]. Hox11-deficient mice show asplenia, indicating that it is essential for genesis of the spleen [12,14].

Subsequently two additional members of the *Hox11* gene family, Ncx/Hox11L1/Enx and Hox11L2, were identified [15]. The *Ncx* gene (Enx, Hox11L1) is specifically expressed in neural crest derived tissues such as dorsal root ganglia, cranial ganglia, sympathetic ganglia, adrenal medulla, and enteric ganglia [16]. Ncx-deficient mice develop megacolon with increased numbers of neuronal cells in enteric ganglia [17,18]. The sequence homology of overall amino acids between Ncx and Hox11 is 61.3%. The homology of homeobox region between them is 86.6%, and especially the helix three sequence that determines a specificity of DNA binding is identical [15]. Although phenotypes of Ncx- and Hox11-deficient mice were distinct from each other, expression of these genes is detected in some overlapping region during embryogenesis and the overlapping regions are intact in each of these deficient mice. These facts suggest that the function of Hox11 family protein is redundant at least in the area where the expression is overlapping.

The Hox11 family proteins belong to a distinct subclass of homeodomain possessing a threonine at position 47 within the third helix. Usually an hydrophobic amino acid, isoleucine or valine, is found at this position [19]. Purified GST–Hox11 homeodomain fusion protein was previously shown to bind to the DNA sequences containing the TAAC or the TAAT core sequence [15]. Moreover, the longer consensus sequence, 5′-GGCGTAAGTGG-3′, was identified using a full length Hox11 recombinant protein [20]. Here we have identified Ncx binding consensus by PCR-mediated selection. Furthermore, we examined transactivation activity of Ncx with a luciferase reporter gene linked to the Ncx binding core sequences and demonstrated that the TAAT core sequence is the target sequence for Ncx. We will discuss the possible mechanism of the transcriptional regulation by Ncx in the context of the Hox11 family protein.

2. Materials and methods

2.1. Expression and purification of GST–Ncx fusion protein

GST–Ncx expression vector was constructed by subcloning a *Not*1 fragment (corresponding to the amino acids 133 to 243 containing homeodomain) of the *Ncx* gene [15] into pGEX-4T2 plasmid (Amersham Pharmacia) in frame. The recombinant plasmid was transformed to *Escherichia coli* strain BL21. Production of GST–Ncx fusion protein was induced by adding 0.1 mM IPTG and the protein was purified by glutathione–Sepharose beads as described in the manufacturer's protocol.

2.2. Antibody production

Polyclonal antibodies against Ncx were raised in rabbits immunized with GST–Ncx recombinant protein using a standard immunization protocol. Immuno-sera were pre-cleared by passage through GST–Sepharose columns followed by affinity purification on antigen columns on which GST–Ncx protein was coupled to activated CH Sepharose (Amersham Pharmacia).

2.3. Identification of the DNA binding site by the selection and amplification binding (SAAB) technique

The 55 or 75 bp oligonucleotide containing a 15 or 35 base randomized internal region (5′-CTGGATCCTAAGATTCCCTG(N)$_{15 \text{ or } 35}$AGGAATTCAGCTTTGAGCCT-3′) was subjected to the SAAB method as described [21]. Briefly, 10 μg of GST–Ncx recombinant protein was incubated with 55 or 75 bp double-stranded random

*Corresponding author. Fax: (81)-43-226 2183.
E-mail: hatano@med.m.chiba-u.ac.jp

PII: S0014-5793(00)01651-3

oligonucleotide in binding buffer (25 mM Tris–HCl, pH 7.5, 50 mM KCl, 10% glycerol, 0.1 mM DTT, 0.1 mM PMSF). The protein–DNA complex was purified by glutathione–Sepharose beads and eluted DNA was amplified by PCR using primers corresponding to the flanking sequences of the random oligonucleotide. The PCR reaction was incubated for 7 min at 94°C for one cycle and 1 min at 94°C, 1 min at 55°C, 1 min at 72°C for 20 cycles. Amplified DNA was mixed with GST–Ncx protein again and the selection was repeated six times. After six rounds of PCR selections, the selected oligonucleotides were subcloned into T easy vector (Promega) and sequenced.

2.4. Electrophoretic mobility shift assay (EMSA)

Double-stranded oligomers were labeled with digoxigenin (DIG) by the DIG 3′ end-labeling kit (Boehringer). 100 fmol of DIG-labeled probe was mixed with 0.1–1 μg of GST–Ncx or GST–Hox11 recombinant protein in total 5 μl of reaction buffer (20 mM HEPES, pH 7.6; 30 mM KCl, 1 mM EDTA, 10 mM $(NH_4)_2SO_4$, 1 mM DTT, 0.2% (w/v) Tween 20, 0.5 μg of poly[(dI-dC)], 0.5 μg of L-lysine) for 15 min at room temperature. The mixture was electrophoresed on a 6% polyacrylamide gel in 0.5×TBE buffer (22.5 mM Tris–HCl, pH 8.5, 28 mM boric acid, 0.7 mM EDTA). The gel was electrotransferred to a nylon membrane (Pall) at 400 mA for 30 min using a semi-dry blotting system (Bio-Rad). The probe was detected by the enhanced chemiluminescent detection system with sheep anti-DIG antibody labeled with alkaline phosphatase (Boehringer). Probes used for this experiment were as follows: a TAAT core; 5′-GTACGGAGTATCCAGCTCCGCCGGTTAATTGGCTCTGG-3′, a TAAG core; 5′-GTACGGAGTATCCAGCTCCGCCGGTTAAGTGGCTCTGG-3′, a TAGT core; 5′-GTACGGAGTATCCAGCTCCGCCGGTTAGTTGGCTCTGG-3′, control (Oct. 1 binding site); 5′-GTACGGAGTATCCAGCTCCGTAGCATGCAAATCCTCTGG-3′.

2.5. Luciferase assay for transcriptional activity

The expression vectors, pAct-Ncx and pAct-Hox11, were constructed by insertion of a full length Ncx or Hox11 cDNA into pHβApr-1 [22] by blunt end ligation, respectively. The luciferase reporter plasmids, pGL3-TAAT and pGL3-TAAG, were constructed by introducing two tandem repeats of the TAAT core (5′-GTACGGAGTATCCAGCTCCGCCGGTTAATTGGCTCTGG-3′) or the TAAG core (5′-GTACGGAGTATCCAGCTCCGCCGGTTAAGTGGCTCTGG-3′) oligonucleotide into the pGL3 luciferase vector (Promega), respectively. For luciferase assay, 0.25 μg of the reporter plasmid and various amounts of the expression vector were transfected into 1×10^5 of C1300 cells by the lipofection method using the Trans Fast kit (Promega) together with 2.5 ng of pRL-SV40 (Promega) as a transfection efficiency control. 3 days later, luciferase activity was measured using the Pikkagene dual (Toyo Ink) with luminometer (Lumat LB9506; Berthold) as described previously [23].

3. Results

3.1. Identification of Ncx binding consensus sequences

In order to determine an optimal DNA binding sequence for Ncx protein, GST–Ncx homeodomain fusion protein was mixed and incubated with double-stranded oligonucleotides containing 15 (N_{15}) or 35 (N_{35}) nucleotides of random core sequences. Specifically bound oligonucleotides were recovered and amplified by PCR. After six sequential rounds of PCR selections, 20 independent clones were randomly picked up from N_{15} and N_{35} selected oligonucleotides respectively and sequenced. Fig. 1 shows an alignment of selected oligonucleotide sequences from N_{15} (A) and N_{35} (B). 10 bp of two related consensus binding sequences, 5′-CGGTAATTGG-3′ (TAAT core) and 5′-CGGTAAGTGG-3′ (TAAG core), were determined.

3.2. Binding of Ncx to the consensus sequences in vitro

Specific interaction between Ncx and the optimal binding DNA sequences was further examined by the EMSA. GST–Hox11 homeodomain fusion protein was used in parallel to

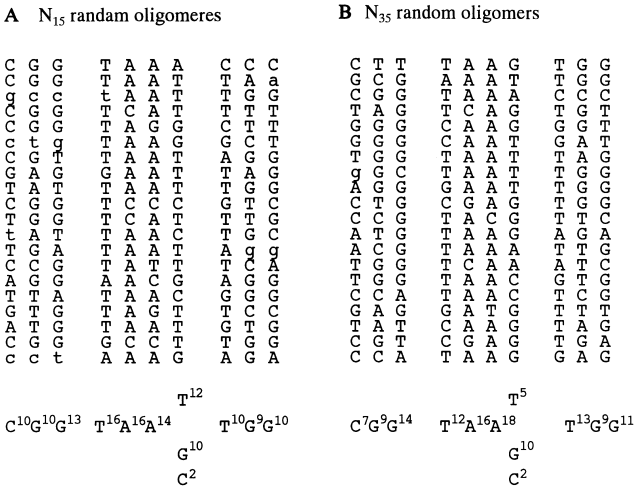


Fig. 1. Alignment of oligonucleotide sequences isolated after six rounds of binding site selection. The best match sequences were determined by Mac Vector II software. The sequences of clones that had a detectable 10 bp consensus sequence derived from N_{15} (A) and N_{35} (B) random oligonucleotide selection are shown. Capital letters indicate the nucleotides corresponding to the original 15 or 35 base region of random sequence. Lowercase letters indicate the region of DNA flanking the 5′ or the 3′ end of the random core element. The nucleotide frequency in each set of clones is indicated below the clone sequences.

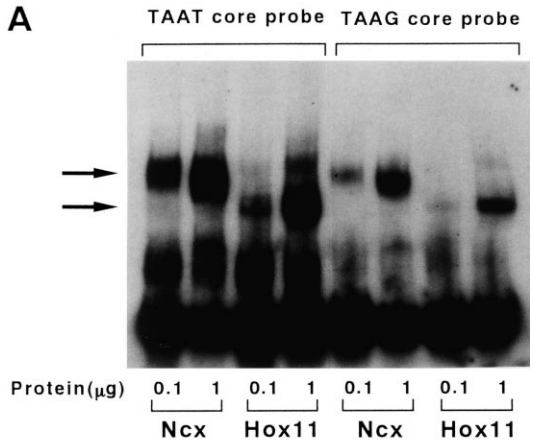
GST–Ncx fusion protein. As shown in Fig. 2, stoichiometric shift bands were observed according to the quantity of recombinant proteins in both Ncx and Hox11 using the TAAT or the TAAG core sequence as a probe. Both Ncx and Hox11 bound the TAAT core in preference to the TAAG core.

To assess specificity and affinity of the protein–DNA interaction, competition assay was performed using unlabeled oligonucleotides. Further studies show that these bindings were specific because they were blocked by unlabeled self oligonucleotides as a competitor but not by mutated oligonucleotides (Fig. 2B and data not shown). Binding of Ncx to the TAAT core probe was completely blocked by 200 fold excess of the cold TAAT core oligonucleotide (Fig. 3A). The TAAG core competitor could also compete Ncx binding to the TAAT probe but less efficiently compared with the TAAT core competitor. On the other hand, when the TAAG probe was used, the TAAT oligonucleotide could compete the binding more efficiently than the TAAG self competitor did (Fig. 3B).

Binding of Hox11 to the TAAT core probe was competed by the TAAT self competitor but not completely by the TAAG competitor (Fig. 4A). On the other hand, the binding to the TAAG core probe was competed by the TAAT core competitor whereas the TAAG self competitor could not so efficiently compete as the TAAT competitor did (Fig. 4B). These results indicate that both Ncx and Hox11 prefer the TAAT core sequence for binding.

3.3. Transcriptional activity of Ncx protein in vivo

We next examined whether Ncx could bind to the selected DNA consensus sequence in vivo and transactivate transcription of a luciferase reporter gene linked to the consensus sequence. Ncx expression constructs and the reporter gene with the TAAT or the TAAG core binding sequence were cotransfected into C1300 neuroblastoma cells. Luciferase activity of



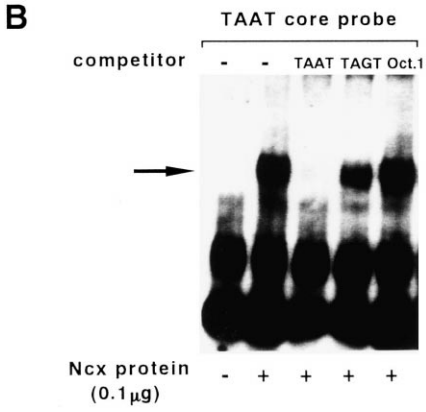


Fig. 2. EMSA with Ncx and Hox11. A: GST fusion proteins containing homeodomain of Ncx or Hox11 were used. The fusion proteins were assayed with two amounts of protein (0.1 and 1 μg) as indicated. The oligonucleotides used for the assays were 5′-CGGTAATTGG-3′(TAAT core) and 5′-CGGTAAGTGG-3′ (TAAG core) as indicated at the top. B: EMSA was performed with reaction mixture without competitor (−) or with 200-fold excess of unlabeled TAAT core, mutated TAAT core (TAGT) or Oct.1 binding sequence as a competitor.
←

the reporter gene linked to the TAAT core sequence was increased in proportion to the amount of the Ncx expression constructs (Fig. 5A). In contrast, when the TAAG core containing reporter plasmid was used, little luciferase activity was detected even with a higher amount of the Ncx expression vector used. This was repeated in another cell line, NIH3T3 (Fig. 5B), indicating that Ncx has stronger transactivation property to the TAAT core consensus binding sequence than to the TAAG core sequence.

Since Hox11 can bind to the same consensus sequences as Ncx protein does, we examined the ability of Hox11 protein to transactivate the same luciferase reporter gene. As shown in Fig. 6, Hox11 has a transactivation ability when the TAAT core consensus sequence was used as a reporter gene. However, when the TAAG reporter plasmid was used, no increase of the luciferase activity was observed. These results show that Ncx and Hox11 are capable of activating transcription in vivo through the TAAT core target sequence.

4. Discussion

We have identified two Ncx binding sequences, 5′-CGGTAATTGG-3′ and 5′-CGGTAAGTGG-3′, by the PCR-based random oligonucleotide selection. These sequences coincided with the Hox11 binding consensus sequences previously reported [15,20]. Since the critical amino acid in the homeodomain for DNA binding is identical between Ncx and Hox11, it is reasonable that both proteins share common consensus sequences for their binding. One of the characteristics of Hox11 family protein is Threonine at position 47 in the homeodomain. Threonine has both a hydrophobic methyl group and a polar hydroxyl group in its side chain and flexibly interacts with different nucleotides. Threonine at this position could potentially bind by hydrogen bonds with the ami-

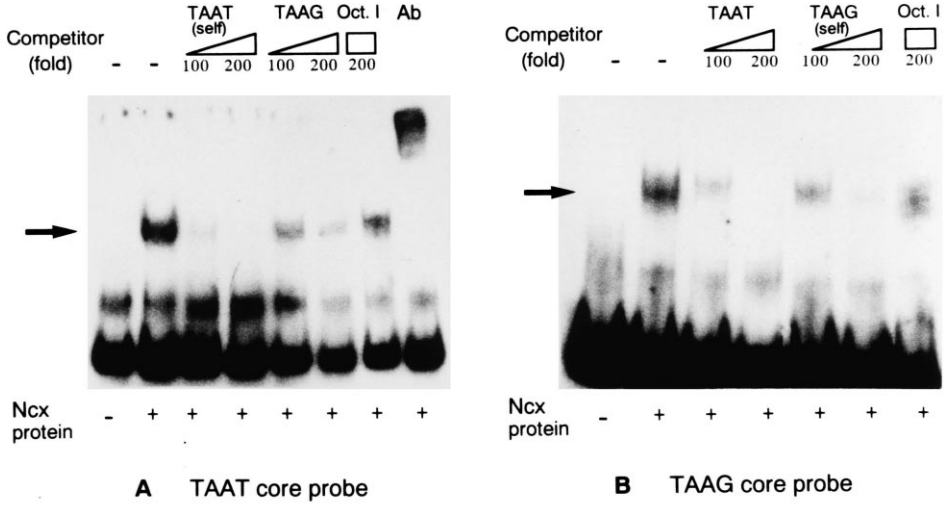


Fig. 3. Competition assays for binding of GST–Ncx fusion protein with the consensus sequence. The TAAT core (A) or the TAAG core (B) probe labeled with DIG was mixed with 0.1 μg of GST–Ncx fusion protein. EMSA was performed with reaction mixture without competitor (−) or with 100- or 200-fold excess of unlabeled TAAT core, TAAG core or Oct.1 as a competitor, or with Ncx specific polyclonal antibodies (Ab) as indicated at the top.

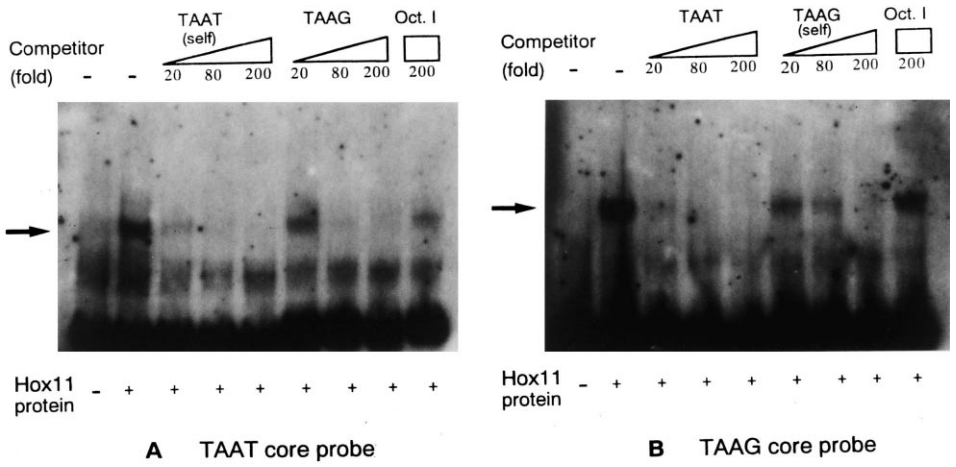


Fig. 4. Competition assays for binding of GST–Hox11 fusion protein with the consensus sequence. The TAAT core (A) or the TAAG core (B) probe labeled with DIG was mixed with 0.5 μg of GST–Hox11 fusion protein. EMSA was performed with reaction mixture without competitor (−) or with 20- 80- or 200-fold excess of unlabeled TAAT core, TAAG core or Oct.1 as a competitor as indicated.

no acid chain of cytosine as well as guanine, or methyl–methyl interaction with a thymine residue [20]. By in vitro selection the TAAT core as well as the TAAG core sequence resulted from random oligonucleotides using GST–Ncx fusion protein. Although the TAAT and the TAAG core sequences were selected in almost equal frequency, affinity of GST–Ncx homeobox protein was stronger to the TAAT core than to the TAAG core. Hox11 also prefers the TAAT core sequence for its binding in our hands.

In other studies the TAAG and the TAAC core were reported to be suitable for the Hox11 consensus sequence [15]. This difference might be due to several reasons. The first was difference in the condition of EMSA assay. We used binding buffer containing 50 mM KCl instead of 30 mM. The second was properties of the recombinant proteins used for the experiments. Though many experiments proved that only the 60 amino acid homeodomain part is sufficient for specific DNA binding, amino acid sequences outside the homeodomain also affect the binding specificity in some cases [24]. The third was the labeling method of the oligonucleotides used for the experiments. We used DIG-labeled oligonucleotide probes instead of ^{32}P and these might influence the protein–DNA interaction.

As these sequences were determined by an in vitro system, we attempted to confirm the transactivation capacity of Ncx

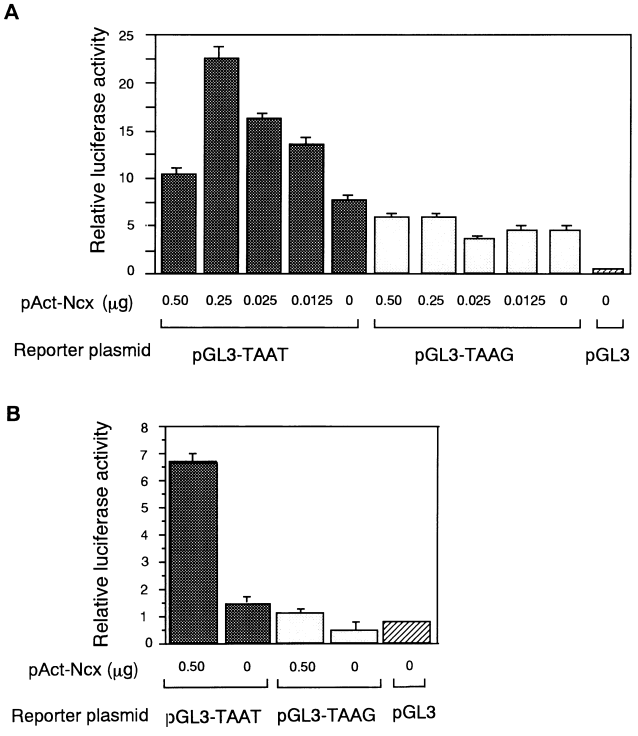


Fig. 5. Luciferase assays from C1300 (A) and NIH 3T3 (B) cells transfected with Ncx expression plasmid. A luciferase reporter plasmid containing two tandem repeats of the TAAT core (pGL3-TAAT) or the TAAG core (pGL3-TAAG) consensus sequence was cotransfected with various amounts of Ncx expression plasmid into C1300 or NIH 3T3 cells. As a control, pGL3 promoter vector (pGL3) was used. The values for the relative luciferase activity were determined from two independent experiments performed in duplicate, where the pGL3 promoter vector was taken as a value of 1.

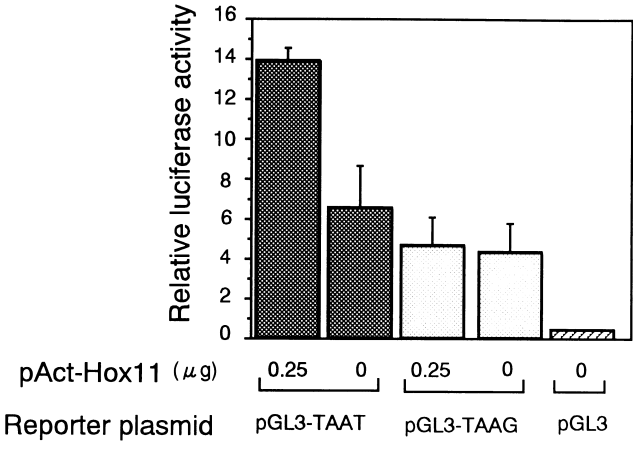


Fig. 6. Luciferase assays from C1300 cells transfected with Hox11 expression plasmid. A luciferase reporter plasmid containing two tandem repeats of the TAAT core (pGL3-TAAT) or the TAAG core (pGL3-TAAG) consensus sequence was cotransfected into C1300 cells with Hox11 expression plasmid. As a control, pGL3 promoter vector (pGL3) was used. The values for the luciferase activity were determined from two independent experiments performed in duplicate, where the pGL3 promoter vector was taken as a value of 1.

and Hox11 by in in vivo transfection using a full length Ncx or Hox11 expression plasmid. Only the TAAT core sequence containing the reporter gene was efficiently transactivated by Ncx. This was also the case with Hox11. Even with a higher amount of the expression plasmids, the TAAG core reporter gene was not activated more than the control level. Thus, in physiological conditions, Ncx and Hox11 prefer to the TAAT core sequence for binding. Interestingly, the TAAT core or TAAG core reporter plasmid alone was significantly transactivated in C1300 cells (Fig. 5A and 6) but not in NIH 3T3 cells (Fig. 5B). Since C1300 cells but not NIH 3T3 express Ncx, this may be due to the endogenous Ncx proteins. Alternatively, other transcription factors which can transactivate the same consensus sequence may be present in C1300 cells but not in NIH 3T3.

The DNA binding capacity of Hox proteins themselves is weak [25]. Although monomeric homeodomain proteins exhibit limited ability to discriminate among different DNA sequences, their specificity is greatly enhanced through cooperative binding of DNA with other DNA binding partners such as Exd in *Drosophila* and Pbx1 in mammalia [26]. The Pbx1 homeodomain containing an additional α helix is larger than the canonical homeodomain. Amino acids of the α helix of Pbx1 interact with the hexapeptide, X(Y/F)PWM(K/R) motif (X indicates a hydrophobic residue), of partner homeodomain proteins and this heterodimer stably binds to DNA [27]. Both Ncx and Hox11 have the hexapeptide motif at 41 amino acids distance in the N-terminal extension of homeodomain and they could potentially heterodimerize to Pbx1. Thus, it is possible that in in vitro selection, monomeric Ncx could not discriminate between the TAAT and the TAAG core. However, in the in vivo situation where Pbx1 protein exists, Ncx and Hox11 can make a heterodimer with Pbx1 and more preferably bind to the TAAT core and transactivate the target sequence. Alternatively, the TAAG core reporter gene may not be transactivated by Ncx or Hox11 under the same situation as the TAAT core reporter gene is. Another cofactor may be required for Ncx or Hox11 to bind to the TAAG core sequence in those cells. Further study is necessary to elucidate a functional significance of the TAAG core consensus binding motif.

Identification of the consensus sequence may be a clue to identify target genes. Although Ncx is expressed in dorsal root ganglia, some cranial nerve ganglia, adrenal medulla and enteric ganglia [16], the phenotype of Ncx-deficient mice was megacolon [17,18]. Similarly, Hox11 is expressed in some cranial nerve ganglia and nucleus, branchial arches and primordium of spleen, but the phenotype of Hox11-deficient mice was asplenia [12,14]. These facts indicate functional redundancy of Hox11 gene family proteins. Since Ncx and Hox11 have overlapping expression at the trigeminus ganglion, these proteins may also differentially and cooperatively regulate putative target genes containing the 5′-CGGTAATTGG-3′ sequence in their regulatory region. Since both Ncx and Hox11 transactivate the reporter gene, these two proteins might be interchangeable in vivo. Analysis of Ncx and Hox11 double knockout mice will reveal the in vivo mechanism of transcriptional regulation by Hox11 family proteins.

Acknowledgements: We thank H. Satake for skillful technical assistance, N. Fujita for secretarial assistance. This work was supported by a Grant-in-Aid for Scientific Research on Priority Areas from the Ministry of Education, Science, Sports and Culture, Japan.

References

[1] Shepherd, J.C., McGinnis, W., Carrasco, A.E., De Robertis, E.M. and Gehring, W.J. (1984) Nature 310, 70–71.
[2] Levine, M. and Hoey, T. (1988) Cell 55, 537–540.
[3] Wolberger, C. (1996) Curr. Opin. Struct. Biol. 6, 62–68.
[4] McGinnis, W. and Krumlauf, R. (1992) Cell 68, 283–302.
[5] Kornberg, T.B. (1993) J. Biol. Chem. 268, 26813–26816.
[6] Nourse, J., Mellentin, J.D., Galili, N., Wilkinson, J., Stabridge, E., Smith, S.D. and Clearly, M.L. (1990) Cell 60, 535–545.
[7] Muragaki, Y., Mundlos, S., Upton, J. and Olsen, B.R. (1996) Science 272, 548–551.
[8] Hatano, M., Roberts, C.W., Minden, M., Crist, W.M. and Korsmeyer, S.J. (1991) Science 253, 79–82.
[9] Dube, I.D., Kamel-Reid, S., Yuan, C.C., Lu, M., Wu, X., Corpus, G., Raimondi, S.C., Crist, W.M., Carroll, A.J., Minowada, J. and Barker, J.B. (1991) Blood 78, 2996–3003.
[10] Kennedy, M.A., Gonzalez-Sarmiento, R., Kees, U.R., Lampert, F., Dear, N. and Rabbitts, T.H. (1991) Proc. Natl. Acad. Sci. USA 88, 8900–8904.
[11] Lu, M., Gong, Z.Y., Shen, W.F. and Ho, A.D. (1991) EMBO J. 10, 2905–2910.
[12] Roberts, C.W., Shutter, J.R. and Korsmeyer, S.J. (1994) Nature 368, 747–749.
[13] Raju, K., Tang, S., Dubé, I.D., Kamel-Reid, S., Bryce, D.M. and Breitman, M.L. (1993) Mech. Dev. 44, 51–64.
[14] Dear, T.N., Colledge, W.H., Carlton, M.B., Lavenir, I., Larson, T., Smith, A.J., Warren, A.J., Evans, M.J., Sofroniew, M.V. and Rabbitts, T.H. (1995) Development 121, 2909–2915.
[15] Dear, T.N., Sanchez-Garcia, I. and Rabbitts, T.H. (1993) Proc. Natl. Acad. Sci. USA 90, 4431–4435.
[16] Hatano, M., Iitsuka, Y., Yamamoto, H., Dezawa, M., Yusa, S., Kohno, Y. and Tokuhisa, T. (1997) Anat. Embryol. 195, 419–425.
[17] Hatano, M., Aoki, T., Dezawa, M., Yusa, S., Iitsuka, Y., Koseki, H., Taniguchi, M. and Tokuhisa, T. (1997) J. Clin. Invest. 100, 795–801.
[18] Shirasawa, S., Yunker, A.M., Roth, K.A., Brown, G.A., Horning, S. and Korsmeyer, S.J. (1997) Nat. Med. 3, 646–650.
[19] Laughon, A. (1991) Biochemistry 30, 11357–11367.
[20] Tang, S. and Breitman, M.L. (1995) Nucleic Acids Res. 23, 1928–1935.
[21] Blackwell, T.K. and Weintraub, H. (1990) Science 250, 1104–1110.
[22] Leavitt, J., Gunning, P., Porreca, P., Ng, S.Y., Lin, C.S. and Kedes, L. (1984) Mol. Cell. Biol. 4, 1961–1969.
[23] Iitsuka, Y., Shimizu, H., Kang, M.M., Sasagawa, K., Sekiya, S., Tokuhisa, T. and Hatano, M. (1999) J. Biol. Chem. 274, 24401–24407.
[24] Hoey, T., Warrior, R., Manak, J. and Levine, M. (1988) Mol. Cell. Biol. 8, 4598–4607.
[25] Chan, S.K., Jaffe, L., Capovilla, M., Botas, J. and Mann, R.S. (1994) Cell 78, 603–615.
[26] Mann, R.S. and Chan, S.K. (1996) Trends Genet. 12, 258–262.
[27] Piper, D.E., Batchelor, A.H., Chang, C.P., Ceary, M.L. and Wolberger, C. (1999) Cell 96, 587–597.